A Multiscale Retinex for Improved Performance in Multispectral Image Classification

B. Thompson[a], Z. Rahman[b], and S. Park[b]

[a]Sandia National Laboratories, P.O. Box 5800, MS 1138, Albuquerque, NM 87185
[b]College of William & Mary, Department of Computer Science, Williamsburg, VA 23187-8795

ABSTRACT

Image preprocessing is useful in helping to identify "spectral response patterns" for certain types of image classification problems. The common artifacts in remotely sensed images are caused by the blurring due to the optics of the image gathering device, illumination variations, and the radiative transfer of the atmosphere. The Multi-Scale Retinex (MSR) image enhancement algorithm that provides dynamic range compression, reduced dependence on lighting conditions, and improved (perceived) spatial resolution has proven to be an effective tool in the correction of image degradations such as those in remote sensing images. In this paper, we measure the improvement in classification accuracy due to the application of the MSR algorithm. We use simulated images generated with different scene irradiance and with known ground truth data. The simulation results show that, despite the degree of image degradation due to changes in atmospheric irradiance, classification error can be substantially reduced by preprocessing the image data with the MSR. Furthermore we show that, similar to the results achieved in previous work, the classification results obtained from the MSR preprocessed images for various scene irradiance are more similar to each other than are the classification results for the original unprocessed images. This is evident in the observed visual quality of the MSR enhanced images even before classification is performed, and in the difference images obtained by comparing image data under different irradiance conditions. We conclude that the application of the MSR algorithm results in improved visual quality and increased spatial variation of multispectral images that is also optimal for certain types of multispectral image classification.

Keywords: image classification, image enhancement, multispectral, retinex, dynamic range compression

1. INTRODUCTION

Various algorithms for image classification using remotely sensed imagery exist in the literature. One way to characterize these algorithms is by the characteristics of the multi-dimensional space in which they operate. For example, many users of remotely sensed imagery utilize spectral signatures to characterize and identify materials in multi-dimensional "spectral" space. The spectral signature of a material may be defined in the solar-reflective region of the electromagnetic spectrum by its reflectance as a function of wavelength, measured at an appropriate spectral resolution. In other spectral regions, signatures of interest are temperature and emissivity (TIR) and surface roughness (radar). The motivation for using remote sensed data for material identification is that different types of materials exhibit different spectral signatures,[1] and so can be distinguished on this basis.

There are fundamental problems with the spectral signature approach that are well documented in the literature. One fundamental problem is that all spectral signatures are unique to the sample *and* to the environment in which they are obtained. Further, the ability to distinguish spectral signatures is often complicated by the natural variability of a material, the spectral quantization of many remote-sensing instruments, and the modulation of the signatures by the atmosphere in the image formation process.[1] Therefore, there is no guarantee that the spectral signatures obtained by the remote sensing system will either be similar to the ones obtained under a different environment, or exhibit measurably different, or even recognizable characteristics.

In recent years, considerable quantities of ground-based (laboratory) data have been accumulated that describe spectral reflectance characteristics of several types of soils and vegetation. However, it is virtually impossible to duplicate natural reflectance variations under laboratory conditions. In addition, the spectral signature of vegetation changes over the seasonal life cycle of plants. Thus the comparison between natural reflectance signatures and laboratory produced signatures becomes even more complicated.

Send correspondence to B.T.: bdthomp@sandia.gov. Co-authors: Z.R.: zrahman@cs.wm.edu; S.P.: park@cs.wm.edu
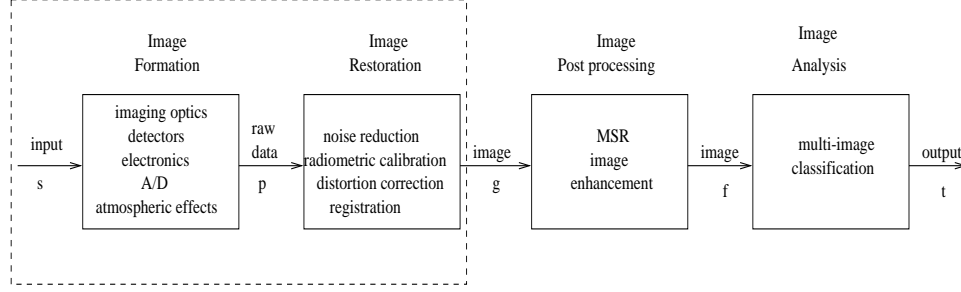
Figure 1. System Model.

As an alternative to classification based on spectral signatures, the multi-dimensional spectral space can be transformed into an application dependent "feature" space that may prove more useful for classification purposes. For example, transformations such as multi-spectral ratios have been used to enhance reflectance differences between types of soil and vegetation, and are used to form "vegetation indices" that aid in classification. Typically, soil and other geological formations exhibit ratios near 1, while vegetation shows a ratio of 4 or more. Other transformations such as scale space filtering have been applied to identify "fingerprints" of certain types of minerals using local points of inflection to characterize absorption characteristics. The success of using these indices has been greatly affected by the lack of suitable methods to account for atmospheric effects on the radiance measured by the remote sensing device.[2]

Figure 1 illustrates the major steps that relate the image acquisition to the image classification process. Regardless of which feature classification method is to be used, the (raw) image data needs to be radiometrically calibrated before it can be used for analysis. That is, the raw data must be converted from sensor DN values to surface reflectance values. Radiometric calibration generally involves several steps: sensor calibration—calculation of gain and offset coefficients that convert sensor DN values to at-sensor radiance values; atmospheric correction—converion of at-sensor radiance values to surface radiance using atmospheric modeling and estimation correction techniques; and, solar and topographic correction—conversion of surface radiance to surface reflectance by correcting for topographic slope and aspect, solar spectral irradiance, solar path transmittance, and down-scattered "skylight" radiance.[1] In addition, detailed information about atmospheric conditions at the time of data acquisition may be required but is generally not available. Parametric atmospheric correction methods can also be used to compensate for atmospheric conditions, but they also require some information about the atmospheric conditions at the time of data acquisition. The success of multi-image classification in the analysis stage is, thus, dependent on the quality of these calibration and correction methods.

In this paper, we approach multi-image classification differently. Instead of applying (parametric) atmospheric correction methods to remote sensed imagery, we compensate for the atmospheric effects by applying the multi-scale retinex (MSR) image enhancement algorithm to the multispectral data prior to classification. The dynamic range compression and color constancy properties of the MSR aid in minimizing the effects of variations in illumination conditions, and the sharpening compensates for the device blurring and atmospheric amplitude modulation.

2. IMAGE PREPROCESSING

There are many factors that contribute to degrade the acquired image. For example, the device signal-to-noise ratio (SNR) and the blurring due to the point spread function (PSF) of its optics, and the quantization artifacts due to the analog-to-digital converters are produced by the image acquisition device. In addition, platform perturbations, atmospheric modulations, and sampling artifacts also degrade the acquired image. Most researchers agree that geometric and radiometric artifacts are the most common cause of image degradations in remotely sensed imagery.[3] With reference to Figure 1 image restoration is an attempt to make the restored image g be geometrically and radiometrically as "close" as possible to the radiant energy characteristics of the original scene s. The *closeness* is measured in some metric space, such as the minimum mean square error (MSRE) space, and the goal of the restoration process is to minimize the MSRE between the restored image and the original scene. Although, generally termed an image enhancement technique, the MSR has proven to be an effective technique for correcting image

Figure 2. 2-d Mondrian.

degradations due to the optical blurring of the image acquisition device, illumination variations, and atmospheric modulation. Thus, the MSR can be used to "restore" the acquired data, without any prior knowledge about the atmospheric conditions at the time of acquisition.

2.1. The Multi-scale Retinex

For all (x, y) pixels in the multi-spectral image G, the multi-scale retinex (MSR)[4,5] can be compactly written as

$$F_j(x, y) = \sum_{n=1}^{N} W_n \cdot \{\log[G_j(x, y)] - \log[G_j(x, y) * H_n(x, y)]\}, \quad j = 1, \cdots, J \tag{1}$$

where J represents the number of spectral bands, N is the number of spatial scales being used, and W_n are the weighting factors for the scales.[6–8] The $H_n(x, y)$ are the surround functions (convolution kernels) given by

$$H_n(x, y) = I_n \exp[-(x^2 + y^2)/\sigma_n^2], \tag{2}$$

where σ_n are the spatial scale parameters that control the extent of the surround function and the I_n are selected so that $\sum \sum H_n(x, y) = 1$. Smaller values of σ_n provide more dynamic range compression, and larger values provide more lightness/color rendition. Each of the expressions within the summation represents a single-scale retinex (SSR).

The MSR combines the dynamic range compression of the small scale retinex with the tonal rendition of the large scale retinex to produce an output which encompasses both. The MSR reduces dependency on lighting conditions/geometry caused by such conditions as obscured foregrounds, and poor lighting caused by atmospheric conditions or defects in artificial illuminants.
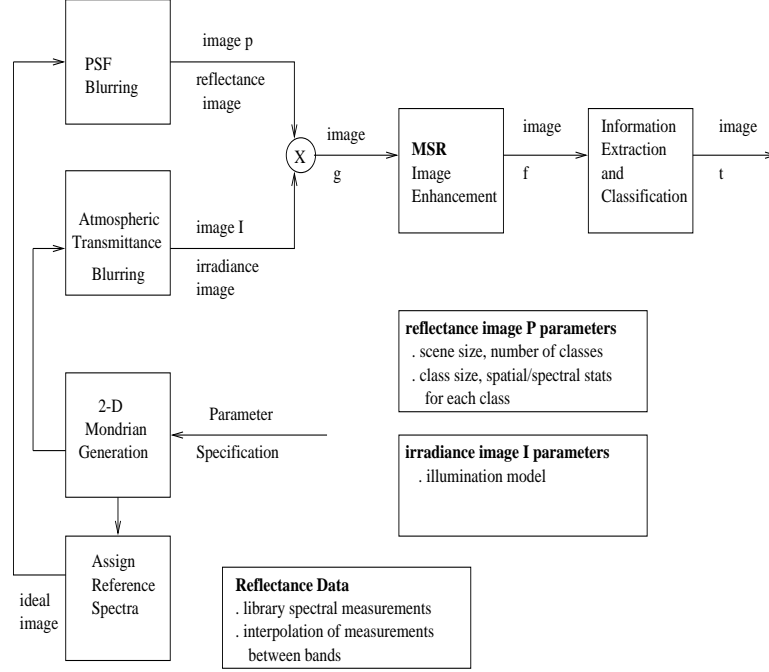
3. SCENE GENERATION

In order to exactly measure the effectiveness of MSR preprocessing on multi-image classification, we use simulated images with exactly known ground truth. The simulated images are created by combining known atmospheric transmittance profiles with scenes with known mean spatial detail and surface reflectance. We use a simple model[9,1] of the radiance field which has the following characteristics:

1. The scene is a two-dimensional Mondrian flat surface divided into patches of uniform reflectance.

2. The effective irradiance (or the atmospheric transmittance) $I(x, y)$ varies slowly and smoothly across the entire scene, and

3. The reflected radiance field $L(x, y)$ is everywhere independent of the viewer's position.

These assumptions permit us to express the radiance field $L(x, y)$ by the simple relationship

$$L(x, y) = \frac{1}{\pi} \rho(x, y) I(x, y), \tag{3}$$

PSF Blurring

image p

reflectance image

Atmospheric Transmittance Blurring

image I

irradiance image

2-D Mondrian Generation

Parameter Specification

Assign Reference Spectra

ideal image

X

image g

MSR Image Enhancement

image f

Information Extraction and Classification

image t

reflectance image P parameters
. scene size, number of classes
. class size, spatial/spectral stats
 for each class

irradiance image I parameters
. illumination model

Reflectance Data
. library spectral measurements
. interpolation of measurements
 between bands

Figure 3. Scene generation system model.

where $\rho(x, y)$ is the Lambertian surface reflectance. The target scene with reflectance $\rho(x, y)$ is represented by the two-dimensional (2-d) Mondrian illustrated in Figure 2. This scene consists of random polygons whose boundaries are Poisson distributed and whose reflectances are distributed according to independent zero-mean Gaussian statistics.

The scene used in the simulation has 256×256 pixels. The reference spectra (See Section 3.1) were subsampled at every $0.05\mu m$ to produce ideal spectra of 42 points. This results in a multi-image with of $256 \times 256 \times 42$ values. Figure 3 illustrates the process by which the 42 band real multi-image is generated. The multi-image is generated by assigning each spatial location in the Mondrian a number corresponding to a specific reference spectra that corresponds to the identification number for a vegetation type.

The atmospheric transmittance function is simulated by significantly blurring a generated Mondrian image. In addition to providing regions of different transmission characteristics, this also simulates the umbra and penumbra profiles of shadows. Two instances of the atmospheric transmittance profile are shown in Figure 4: the dark regions represent absorption in the atmosphere, or clouds; and the light regions represent transmittance, or absence of clouds. The first profile, atrans1, has two moderately sized regions of low transmittance ($< 30\%$) at the lower left and right corners of the image. For the majority of the scene the transmittance is about 45%, with some higher transmittance areas located at the bottom center portion of the image which have a transmittance of about 80%. The second profile, atrans2, has on average a transmittance of about 65%. As shown in Figure 3, after the ideal scene image m and atmospheric transmittance profile m' have been generated, the simulated multi-image g is obtained by doing a pixel by pixel multiplication of m' and m,

$$g(x, y) = m(x, y) \cdot m'(x, y)$$

Finally, the real image g is processed with the MSR algorithm to create the processed image f.

In order to analyze the data, fidelity metrics are computed for the ideal, real, and the MSR processed images. Because we have "ground truth" data, these fidelity metrics can be applied and the results compared to the ground truth. The fidelity analysis is based on two metrics that measure the accuracy and consistency of the results as they are affected by the application of the two atmospheric transmittance profiles. A *mean squared-error* metric is used to measure the similarity between the two original images, and the similarity between the two real retinex images for each atmospheric transmittance profile. A *sensitivity* metric is used to measure classification consistency and
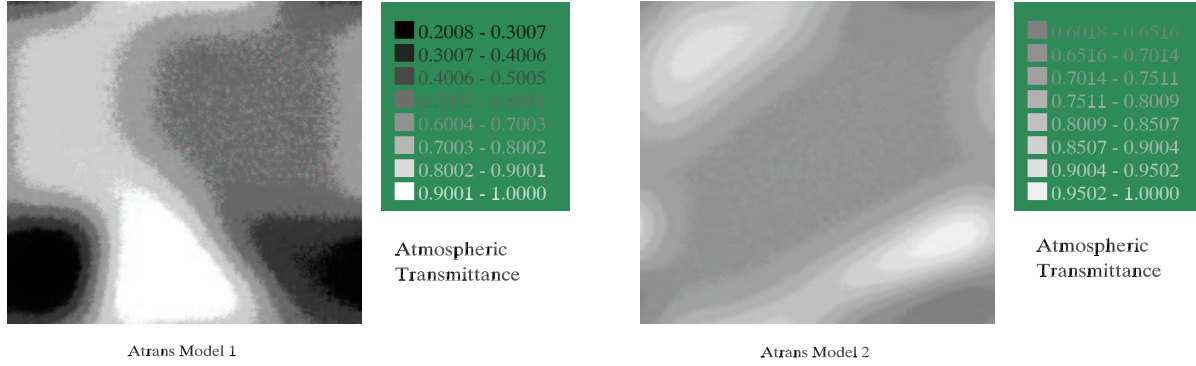
Figure 4. Transmittance Models.

accuracy for the real original and real MSR processed images using the real ground truth and the MSR ground truth as the basis of comparison.

3.1. Agronomic Data

The reference spectra used to create the ideal Mondrian scene were taken from an AVIRIS (Airborne Visual and Infrared Imaging Spectrometer) dataset for the San Luis Valley in Colorado.[10]   Reference spectra plots for 9 representative species of vegetation are used in the scene (Figure 5), of which 8 are used in this experiment. The study used farmland reference spectra representing potatoes, alfalfa, barley, oat hay, canola, and open fields containing chico.[10]   Based on the analysis provided by the USGS study by Clark et al,[10] the alfalfa, canola, oat hay, and nugget potato spectra showed the plants to be green and healthy. The barley had lost all of its chlorophyll signature. The norkotah potatoes were not being irrigated as they were about to be harvested, and consequently showed weak chlorophyll and cellulose absorptions, with soil (clay) absorptions from exposed soil. These potatoes were also being sprayed with a defoliant, so they showed decreased chlorophyll absorption, and a shift of the red edge of the absorption spectrum to shorter wavelengths. The chico and pasture spectra showed combinations of chlorophyll and cellulose (dry vegetation) absorptions. There was rain in the valley in the few days before the data acquisition flight so the chico and pasture did not show much water deprivation stress—being native plants they are hardy and can also withstand more reduced precipitation compared to the crops. The bare field calibration spectrum is from a sample measured on a laboratory spectrometer; all others are averages of several spectra extracted from the AVIRIS data.

3.2. MSR Preprocessing

Figures 6 and 7 show RGB and linear contrast stretched (LCS) composites of the real original and MSR processed scene. In Figure 6(l), the original scene has a reflectance of at most 60, thus making it difficult to compare it to the MSR processed image of Figure 6(l). However, in the LCS versions of the images we are able to compare features in both images. The most striking observation between the two images is that the MSR enhances details between class borders and within the class regions so that the sharpness of features in the image distinguishes it over that of the original image. We do note the appearance of edge artifacts within the borders of the regions for the MSR image. These edge artifacts are caused by Mach band undershoots and overshoots displayed as dark boundaries around the border of certain regions.[11]

4. DISCUSSION

Figure 8 shows the real original and MSR images created with the two atmospheric transmittance profiles shown in Figures 4a and 4b. The original images have been linearly stretched so that subtle differences between them and the MSR image can be compared. Recall that atrans1 (Figure 4(l)) has an average transmittance of about 45%, and atrans2 (Figure 4(r)) an average transmittance of 65%. Comparing the effects of the transmittance profiles on the original and MSR images, it is evident that the MSR images appear visually consistent for both the models. However, the original images show the effects of the transmittance models as various dark reflectance areas in the images. For instance, the affect of the low transmittance areas is easily seen in the bottom left and right portion of
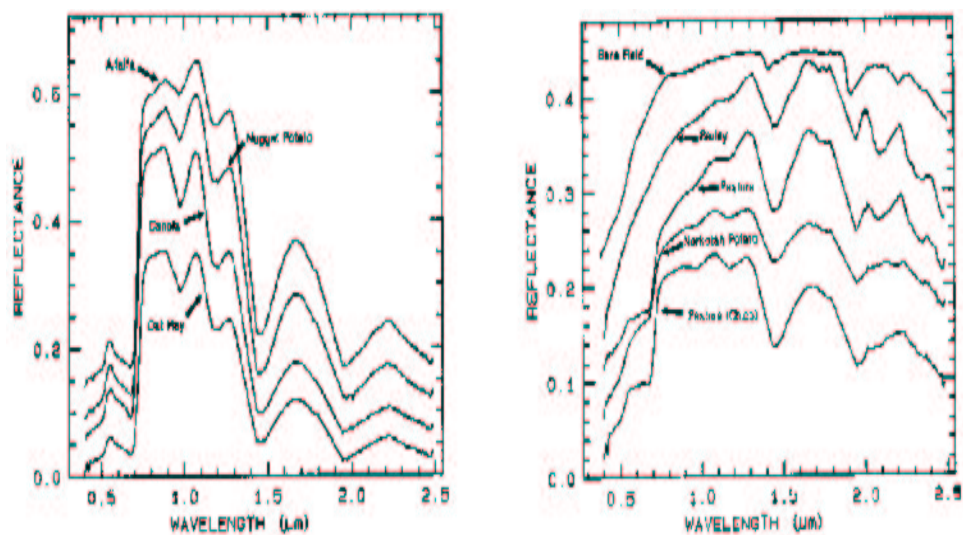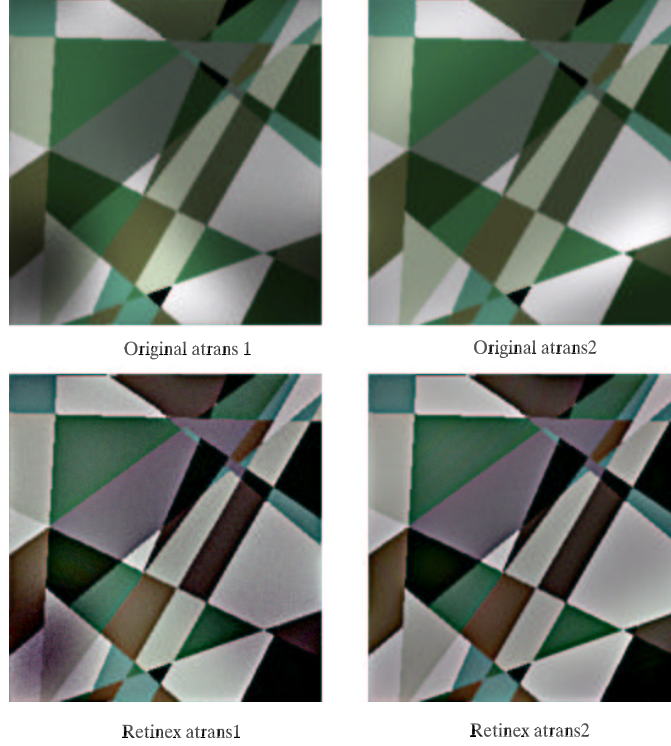
Figure 5. Reference spectra of vegetation and soil types. (Source: USGS Speclab.[10])

Figure 6. real 2-d Mondrian scene: (l) original (r) linear contrast stretched version.
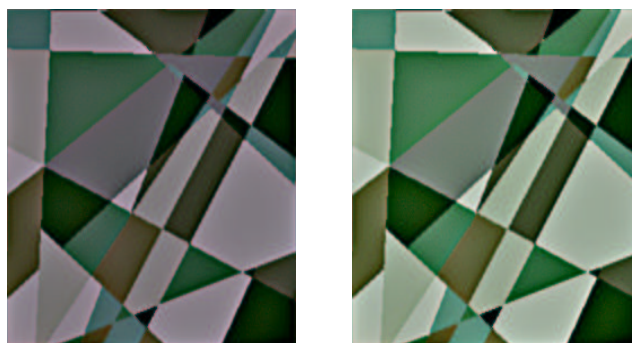
Figure 7. real 2-d Mondrian scene: (l) MSR processed (r) linear contrast stretched version.

Original atrans 1

Original atrans2

Retinex atrans1

Retinex atrans2

Figure 8. Effect of atmospheric models on original and MSR processed images.

the original atrans1 image. Very bright areas of reflectance in the original atrans2 image are also discernible in areas where the transmittance was the highest. In the MSR atrans1 image there is some slight darkening of reflectance in the lower left corner of the processed image, however, overall the MSR images are more similar to each other than are the unprocessed images. We also observe that the boundaries between regions in the MSR images are more clear and are in greater detail than the original. The previously identified edge artifacts are, however, evident at the transition between dark and bright areas in the processed images.
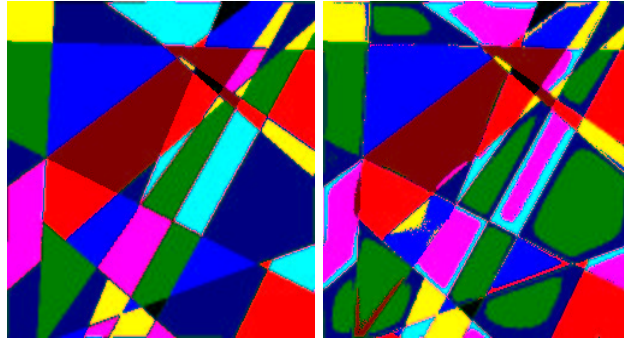
5. CLASSIFICATION

We used vector quantization (VQ) to perform unsupervised classification on the multi-spectral image. The only user specified parameter is the number of classes K. For classification, a 9 band subset of features were chosen from the original 42 band image.

To cluster the images we used VQ along with a splitting method to define the spectral signatures.[12] The splitting algorithm used to generate the trained codebook, splits each training set codebook vector using the best perturbation factor for that dataset. The preferred perturbation factor is the one that generates the smallest MSE for the input training set. The algorithm is designed to produce cluster means for a specific codebook size. As is the case for most classification methods, the performance depends on the quality of the set of spectral means used to discriminate classes in the image. For this analysis, we did not focus on methods to obtain spectral means, but compared the relative accuracy of the spectral means obtained by the VQ to signatures derived from the training areas

Because we have a ground truth map of our ideal classification, the training set will be selected from regions in the image shown in Figure 2. The training set vectors were input into the splitting algorithm and the trained codebook vectors were generated. In the testing stage, the images were classified with the trained codebook vectors, using a MSE VQ clustering algorithm. The resulting test codebook vectors were used as candidate spectral means to identify each vegetation species.
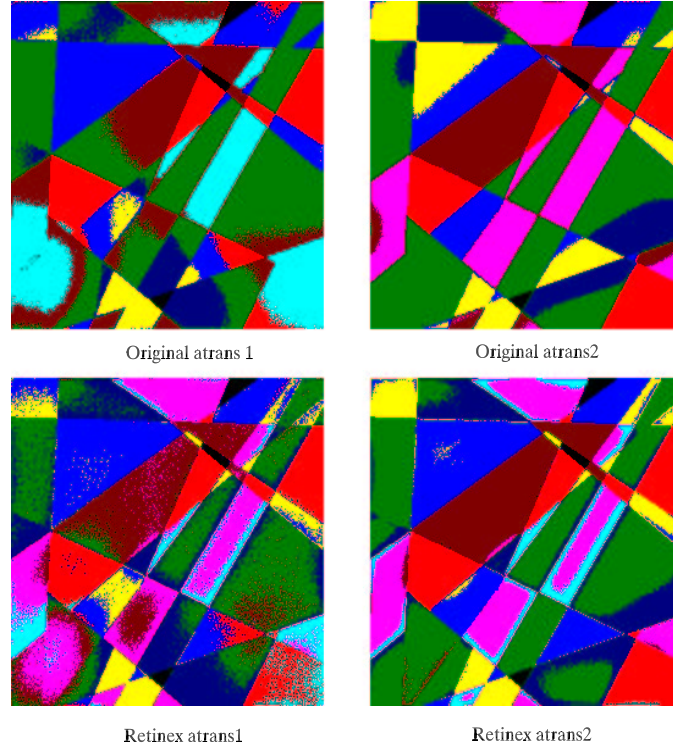
Figure 9. Classification results for original and MSR — Atmospheric Models Applied.

Figure 10. Classification results for original and MSR — No Atmospheric Models Applied.

Figure 9 shows the classification results for the original and MSR images for the two transmittance profiles. The classification results for the MSR processed images are more consistent with each other than the results obtained with the original images. However, there were problems separating certain class pairs such as pasture (c) and potato (nh), and barley and bare field. The pasture (c) spectra is not identified in the atrans 2 model image for the original. Both the original and the MSR processed atrans 1 images have classification errors resulting from the affect of the low transmittance area in the lower left portion of the images.

Tables 1 and 2 show the classification sensitivity results for the original and MSR image for different ground truth images. The original and MSR ground truth images refer to the images shown in Figure 10. The sensitivity measurements are listed in terms of the percentage of pixels correctly classified and the actual number of pixels correctly classified out of the total 65536 (256 × 256) pixels that belong to the classified image. As can be seen in Table 1, the MSR processed image for the atrans1 transmittance profile provided better classification sensitivity. When used with the real *original* ground truth image, the MSR produced sensitivity results slightly better, 2.5%,

than those produced by using the original atrans1 data. However original image for the atrans 1 model. However, when the MSR processed original data was used for training, the classification accuracy for the MSR processed image was substantially better, $\sim 15\%$, than the original. An even greater separation in results is seen between the original and MSR images for the atrans2 transmittance profile. When the MSR processed data is used for training, the MSR classified image is approximately 23% better than the original classification. Even when the original unprocessed data is used for training, the MSR classification is about 10% better.

Table 1. Sensitivity (# pixels out of 65536)

Ground Truth	Image	Model 1	
		% correct	# pixels correct
Real Original	Original	57.56	37722
Real Original	MSR	60.99	39970
Real MSR	MSR	70.63	46288

Table 2. Sensitivity (# pixels out of 65536)

Ground Truth	Image	Model 2	
		% correct	# pixels correct
Real Original	Original	54.32	35599
Real Original	MSR	74.34	48719
Real MSR	MSR	86.86	56924

Figure 11 shows an RGB composite of the quantized images. The original images have been linearly stretched so that subtle differences between them and the MSR image can be compared. Comparing the effects of the atmospheric models on the original and MSR images, we see that the MSR images appear visually consistent between both models which is consistent with the results obtained using the 42 band multispectral image results discussed earlier in this experiment. However, the effects of the transmittance models on the original images is more apparent in these quantized images. The low transmittance areas are sharply contrasted with the high transmittance regions. In the MSR atrans1 image we do observe the same darkening of reflectance in the lower left corner of the image which corresponds to the same low transmittance in the atrans1 model in the same area. We observe the same edge artifacts effects around the borders of regions in the MSR quantized images as seen in the 42 band MSR image. For this experiment, the advantage of generating classification and quantized images simultaneously is that the classified image may provide an indication of how closely the quantized image will match the original data.

In Figure 13 we illustrate the results of using the squared-error difference metric to compare the original and MSR images for each atmospheric transmittance model. We conclude from the difference images that the consistency observation between the MSR images for the difference transmittance models is again confirmed. These results are similar to the results shown in Figure 12. However, we observe more high difference regions in Figure 13 for the original image, than in Figure 12.

6. CONCLUSIONS

Although image enhancement is typically applied to improve the visual quality of multispectral images, in this experiment we have given quantitative evidence that the application of the MSR algorithm restores images that are degraded by atmospheric transmittance effects, and improves the results of multispectral image classification. Because the MSR algorithm was applied before clustering, the classification algorithm generated candidate spectra that were better separated in reflectance for the MSR images than the spectra generated for the original images. Furthermore, the MSR candidate spectra maintained separability and high reflectance values regardless of the atmospheric transmittance models applied. This leads us to conclude that the application of the MSR algorithm produces (approximately) illuminant invariant spectral signature images. Except for class regions in which edge artifacts produced incorrect classifications around region boundaries, the classification results and the difference measurement results show a consistency between MSR images that is not evident in the classifications based on the original images.
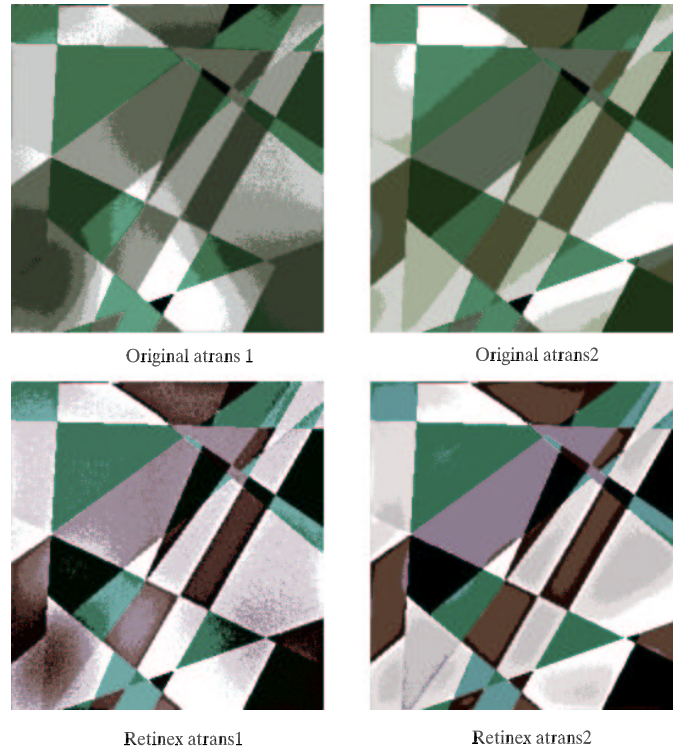
Figure 11. Effect of atmospheric models on VQ original and MSR images.
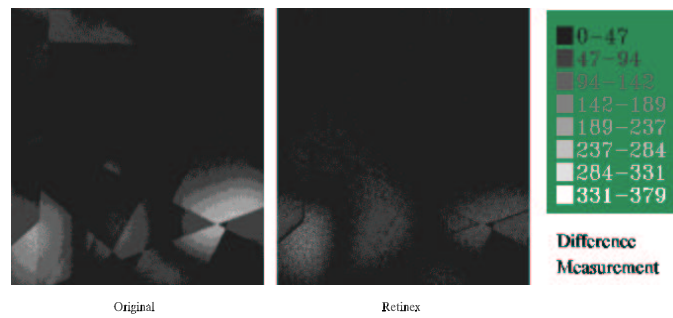
Figure 12. Difference images for the (l) original and (r) MSR images.
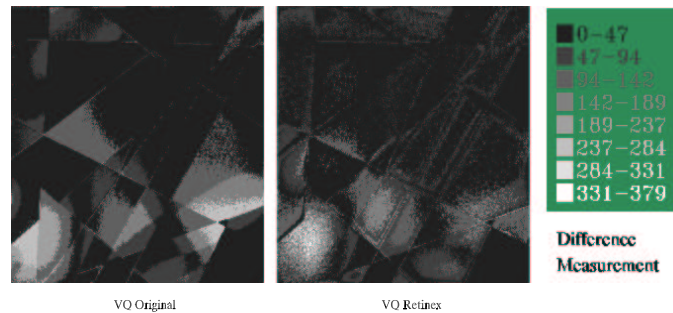
Figure 13. VQ Difference images for the (l) original and (r) MSR images.

REFERENCES

1. R. A. Schowengerdt, *Remote Sensing: Models and Methods for Image Processing*, Academic Press, 1997.
2. R. Jackson, P. Slater, and P. J. Pinter, "Discrimination of growth and water stress in wheat by various vegetation indices through clear and turbid atmospheres," *Remote Sensing of Environment* **13**, pp. 187–208, 1983.
3. J. R. Jensen, *Introductory Digital Image Processing: A Remote Sensing Perspective*, Prentice-Hall, Inc., 1996.
4. E. H. Land, "The retinex theory of color vision," *Scientific American* **237**, pp. 108–129, December 1977.
5. M. Olshaker, *The Instant Image: Edwin Land and the Polaroid Experience*, Stein and Day, Scarborough House, Briarcliff Manor, N.Y., 1978.
6. Z. Rahman, D. J. Jobson, and G. A. Woodell, "Multiscale retinex for color image enhancement," in *IEEE International Conference on Image Processing*, vol. 3, pp. 1003–1006, 1996.
7. D. J. Jobson, Z. Rahman, and G. A. Woodell, "Properties and performance of a center/surround retinex," *IEEE Transactions on Image Processing* **6**, pp. 451–462, March 1997.
8. D. J. Jobson, Z. Rahman, and G. A. Woodell, "A multiscale retinex for bridging the gap between color images and the human observation of scenes," *IEEE Transactions on Image Processing* **6**, pp. 965–976, July 1997.
9. F. O. Huck, C. L. Fales, R. E. Davis, and R. Alter-Gartenberg, "Visual communication with retinex coding," *Applied Optics* **39**, April 2000.
10. R. N. Clark, V. King, and G. A. Swayze, "Initial vegetation species and senescence/stress indicator mapping in the San Luis Valley, Colorado using imaging spectrometer data," 1995. http://speclab.cr.usgs.gov/PAPERS.veg1/vegispc2.html.
11. R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, Addison-Wesley, 3rd ed., April 1992.
12. K. Sayood, *Introduction to Data Compression*, Morgan-Kaufmann, 1995.